Educational Evaluation and Policy Analysis
http://eepa.aera.net

Keeping At-Risk Students in School : A Systematic Review of College Retention Programs

Jeffrey C. Valentine, Amy S. Hirschy, Christine D. Bremer, Walter Novillo, Marisa Castellano and Aaron Banister

EDUCATIONAL EVALUATION AND POLICY ANALYSIS 2011 33: 214
DOI: 10.3102/0162373711398126

The online version of this article can be found at:
http://epa.sagepub.com/content/33/2/214

Published on behalf of

American Educational Research Association

and

http://www.sagepublications.com

Educational Evaluation and Policy Analysis
June 2011, Vol. 33, No. 2, pp. 214–234
DOI: 10.3102/0162373711398126

Keeping At-Risk Students in School: A Systematic Review of College Retention Programs

Jeffrey C. Valentine
Amy S. Hirschy
University of Louisville

Christine D. Bremer
Walter Novillo
University of Minnesota

Marisa Castellano
Aaron Banister
University of Louisville

The wage premium for college graduates is substantial. Far fewer of these benefits accrue to students who complete some college (i.e., those who do not persist to graduation), and, partly for this reason, colleges often adopt programs aimed at helping to keep at-risk students in school. This article reports on a systematic review and meta-analysis of studies of such retention programs. The studies suggest small but potentially important effects on short-term retention rates and grades earned by program participants. Studies of more comprehensive interventions using relatively more appropriate comparison groups suggested more effective results than did studies that used weaker interventions, relatively less appropriate comparison groups, or both. Even the best studies included in this review are methodologically suspect and as such do not provide a very strong basis for making policy recommendations. From a public policy perspective, this review points to the need for more investment in rigorous studies that investigate, at a finer level of detail, the specific aspects of programs that are associated with program success. Rigorous studies are also needed that investigate the interaction between programs and student characteristics to determine what types of programs are most effective for which students.

Keywords: *retention, dropout prevention, meta-analysis*

The work reported herein was supported under the National Research Center for Career and Technical Education, PR/Award (No. V051A070003), as administered by the Office of Vocational and Adult Education, U.S. Department of Education. However, the contents do not necessarily represent the positions or policies of the Office of Vocational and Adult Education or the U.S. Department of Education, and you should not assume endorsement by the federal government. We thank Jay Noell, Jim Stone, and three anonymous reviewers for the National Research Center for Career and Technical Education for their helpful suggestions and Kirsten E. Sundell for editing previous versions of this article.

Students plan to participate in postsecondary education for a variety of reasons. However, the great majority of them are probably motivated, at least in part, by the economic returns associated with postsecondary education. In debates over who should pay for and who benefits from investment in postsecondary education, most agree that education beyond high school is both a private and a public good. Although a higher level of education is not the only factor involved in earning higher wages (Kemple, 2008), there are, on average, advantageous economic returns for increased levels of education: Returns to baccalaureate degrees surpass returns to associate degrees, and those with an occupational certificate (equivalent to two semesters of full-time study) have higher earnings than those with just some college but no degree (Grubb, 2002; Marcotte, Bailey, Borkoski, & Kienzl, 2005). As a public good, it seems clear that a nation's economic status depends in part on the quality and quantity of postsecondary education available (Barton, 2008; Paulsen & St. John, 2002). Worker groups with more education, including career-related education (Kemple, 2008), tend to have higher employment rates (Krolik, 2004; Prince & Jenkins, 2005), and an educated citizenry makes fewer demands on publicly funded resources, such as social services, welfare, and corrections (National Center for Public Policy and Higher Education [NCPPHE], 2004). People with higher levels of income also generate more tax revenue and economic activity (Barrow & Rouse, 2005; Barton, 2008).

The majority of positions that pay wages or salaries high enough to support a family—and almost two thirds of all jobs—require skills associated with at least some education beyond high school (Carnevale & Derochers, 2003). In fact, although high school grades and behaviors are associated with long-term employment and earnings outcomes (Rosenbaum, DeLuca, Miller, & Roy, 1999), many researchers would agree with Rosenbaum's (2001) assertion that high school records (i.e., grades, attendance, test scores) have little relationship to employment or earnings immediately after high school. Employers place little stock in such records; furthermore, recent high school graduates are rarely hired for demanding careers—they typically are hired for entry-level jobs instead. Many students thus enroll in postsecondary education (e.g., university, community college, or programs designed to lead to qualification for skilled jobs) in an attempt to improve their employment prospects.

Mere postsecondary enrollment is insufficient, however. When postsecondary education totals less than a year, earnings increases are negligible (Grubb, 2002). Understanding the educational transitions that students must navigate into and through postsecondary systems is therefore critical to improving opportunities for all students and for disadvantaged students in particular. Dropout can occur at any of several stages in the process of moving from secondary into and through postsecondary education. For example, students may not have been adequately prepared for the academic requirements of postsecondary education. They may have difficulty balancing employment, family, and education commitments; they may believe that they are not suited for college; or they may feel out of place and find it difficult to make friends and find social support in the college setting. Transition and retention programs aim to help students succeed in the face of such challenges and attain their educational goals.

This review begins with a brief overview of the theoretical frameworks that address college choice and postsecondary student departure processes, followed by a literature review of programs that are designed to increase persistence rates among at-risk college students. The Method section describes the search strategy used to locate studies and provides specific information on how we extracted data, evaluated studies for inclusion, and synthesized the results of the studies that were found. In the Discussion section, we highlight the most important findings, discuss the need for funding and conducting additional rigorous research on these programs, and suggest characteristics of research that are important for framing public policy.

Theoretical Frameworks

Researchers and practitioners often use theoretical formulations to better understand how their efforts can be more effective. To understand and address the challenges related to secondary to postsecondary transition, several theoretical

perspectives are reviewed here. The integration of multiple theoretical perspectives is currently a strong trend in the study of college choice and student persistence phenomena. Researchers have drawn from economic (e.g., Cabrera, Nora, & Castaneda, 1993; St. John, Paulsen, & Carter, 2005), sociological (e.g., Paulsen & St. John, 2002; Perna & Titus, 2005), organizational (e.g., Bean, 1980, 1983), and psychological (e.g., Astin, 1984; Eaton & Bean, 1995) theoretical and conceptual models to understand these ill-structured problems (Braxton & Mundy, 2001–2002). Below we review some of these considerations.

College Choice Process

A student's choice to attend college is described as a "complex, multi-stage process during which an individual develops aspirations to continue formal education beyond high school, followed later by a decision to attend a specific college, university, or institution of advanced vocational training" (Hossler, Braxton, & Coopersmith, 1989, p. 234). Precollege students cycle through three stages: predisposition, search, and choice (Hossler & Gallagher, 1987). Throughout their primary, middle, and high school experiences, students determine whether to attend postsecondary education, search for information about institutions and financial aid, narrow the options, and then decide on an institution (Hossler et al., 1989; McDonough, 2004).

More recent research approaches the college choice process with an integrated conceptual model, drawing from economics of human capital models and sociological theories of cultural and social capital (Perna, 2006). Human capital theory predicts that an individual's investments (in this case, investing time and financial resources in education and training) to increase his or her capabilities will pay off in improved financial status and quality of life (Becker, 1993). Sociological theories of cultural, social, and economic capital recognize that resources vary among socioeconomic strata of society, and these resources, or capital, take many forms. Cultural capital refers to the system of characteristics that defines an individual's class status. Resources such as language skills, cultural knowledge, and mannerisms derived in part from one's parents are examples of cultural capital (Bourdieu, 1986; Bourdieu & Passeron, 1977). Social capital theory focuses on how individuals acquire human, cultural, and other forms of capital through their memberships in social networks (Coleman, 1988). For example, through their relationships with teachers, advisors, and peers, high school students may or may not learn about a wide array of postsecondary options, careers, or financial aid resources.

St. John, Paulsen, and Starkey (1996) explored the linkage or "nexus" between a student's college choice and persistence processes (p. 175), finding that when a student's finances influenced the choice of institution, they later interacted with academic and social experiences and the student's decision to persist in college. Conflicting findings about the role of finances in the college choice and persistence processes reveal complexity (Nora, Barlow, & Crisp, 2006). For example, student perceptions of affordability and types of financial aid sometimes vary by socioeconomic and cultural backgrounds (Nora, Barlow, & Crisp, 2006; Schuh, 2004). St. John, Cabrera, Nora, and Asker (2000); St. John, Asker, and Hu (2001); and Schuh (2004) provide useful reviews of the research on student finances in the college choice and persistence processes.

Postsecondary Student Departure

Tinto's Interactionalist theory (1975, 1982, 1986, 1987, 1993) primarily addresses voluntary student departure decisions within postsecondary institutions. Tinto (1975) postulated that students enter college with a variety of personal, family, and academic characteristics and skills. These are subsequently affected by the interactions of the student with the college community. Importantly, these individual characteristics influence decisions to stay in or leave college. Central to Tinto's theory is the degree to which a student becomes integrated into the academic and social realms of the institution. Academic and social integration influence a student's subsequent commitment to the institution and to the personal goal of college graduation. Finally, Tinto postulates that the greater the levels of institutional commitment and commitment to the goal of college graduation, the more likely the individual will persist in college.

Though considered paradigmatic in stature (Braxton, Hirschy, and McClendon, 2004), researchers have tested and critiqued Tinto's theory. Braxton, Sullivan, and Johnson (1997) reviewed empirical studies based on Tinto's theory, concluding that it lacks internal consistency and needs revision. In Braxton's *Reworking the Departure Puzzle* (2000), leading scholars (including Tinto) aimed to update the theory to reflect current demographical, cultural, and institutional perspectives. For example, Rendón, Jalomo, and Nora (2000) note that students with varying ethnic and gender backgrounds may experience the social and academic context differently, yet researchers and practitioners often assume that Tinto's assumptions can be applied to all students in the same way. If Tinto's interactionalist model is based on an acculturation/assimilation perspective, then minority students may be pressured to separate from their cultural communities to successfully integrate into the college environment. Such separation as a rite of passage may hold harmful consequences for racial and ethnic minorities (Tierney, 1992). Hurtado and Carter (1997) add that the nature of the integration concept and the timing of when it is measured in a student's academic career may affect research results. Rendón et al. (2000) recommend that future research include qualitative methods that allow a definition of social integration to emerge from minority students' perspective, as the current paradigm is limited.

Tinto's (1982, 1986, 1987, 1993) revisions to his theory addressed the importance of financial resources within the set of background characteristics with which a student enters a postsecondary institution and acknowledged the role communities external to the institution (e.g., family, work, and neighborhood) play in students' departure decisions. Similarly, Bean and Metzner's (1985) model builds on Tinto's but emphasized the importance of external influences on the persistence of nontraditional students, such as those at community colleges. More recently, Braxton et al. (2004) theory of student departure in commuter institutions gives greater importance to both the internal campus atmosphere (e.g., academic communities and institutional environment) and students' life circumstances away from campus.

Many programs designed to increase persistence among at-risk college students build implicitly or explicitly on Tinto's (1993) theory. For example, orientation programs often introduce students to campus academic resources (e.g., tutoring programs, librarian assistance). Similarly, faculty mentoring programs are aimed in part at increasing the academic integration of students. Institutions also attempt to increase social integration by offering and supporting activities that serve to connect students to each other. Activities of this sort include Greek life and intramural sports. Finally, interventions such as learning communities may foster both academic and social integration. To date there has yet to be a state-of-the-art, systematic review on the effectiveness of interventions designed to increase persistence rates among at-risk college students. As such, we set out to conduct a systematic review and, if appropriate, a meta-analysis (or several meta-analyses) of studies that attempt to determine the effectiveness of these programs. Note that the distinction we draw between a systematic review and a meta-analysis is intentional. We define a systematic review as a systematic and transparent approach to the collection and evaluation of a literature on a specific research question and meta-analysis as the quantitative analysis of the results of multiple studies. A systematic review need not include a meta-analysis (e.g., one can systematically and rigorously look for literature on a research question and not find any), nor must a meta-analysis be based on studies located through a systematic reviewing process (in most cases, though, a meta-analysis should be based on a systematic review) (Valentine, Pigott, and Rothstein, 2010). In the sections that follow, we detail (a) the methods of the systematic review and meta-analysis, (b) the results of the meta-analysis, (c) the methods and results of one study that met our inclusion criteria but could not be included in the meta-analysis, (d) the limitations of the review, and (e) suggestions for both policymakers and future research.

Method

Inclusion Criteria

To be included in this review, a study had to meet two criteria. Because we were interested in programs for students who could be considered at-risk, our first criterion was that the sample had

to be composed primarily of students who fit a relatively broad definition of at-risk status (including for example academically underprepared and economically disadvantaged students). In addition, a study had to provide a quantitative evaluation of a formal program or intervention that was aimed at keeping at-risk college students in school. Ideally, the evaluation would involve random assignment to conditions in the context of an experiment (e.g., of students to either a prevention program or to a wait-list control), but we anticipated that these would be rare; as such, we required only that a study include a local comparison group. Our rationale for this decision is that it is becoming increasingly clear that having a local comparison group improves efforts to equate groups (see Cook, Shadish, & Wong, 2008). Functionally, the requirement of a local comparison group meant that we included studies that, for example, exploited existing data resources (such as a university's database) and using dummy coding to identify students who attended or did not attend a program for at-risk students (a number of the studies in our review used this analysis strategy). However, we would not have included a national survey that employed a similar analysis strategy. See the section "Quality Assessment of Included Studies" for details of our quality assessment plan.

Literature Search

We used several strategies to find relevant literature, as recommended in the literature on systematic reviewing (see Rothstein, Turner, & Lavenberg, 2004). First, we conducted searches of the electronic databases PsycINFO and ERIC. We used a series of categories, connecting terms within categories by an OR statement, and across categories by an AND statement. We searched for terms suggesting a college transition program (e.g., transition OR college), an empirical study (e.g., outcome OR results OR compare*), a program (e.g., program OR intervention), and student risk status (e.g., at-risk OR dropout OR underprepared). The titles and abstracts from citations identified via the database search were examined by at least two of the authors working independently. We obtained full copies of a study if, after discussion, both individuals agreed that it might meet our inclusion criteria.

In addition, we identified 124 documents that could, broadly speaking, be conceptualized as literature reviews of transition programs. For each of these reviews, we examined the reference sections for citations that appeared to meet our inclusion criteria and attempted to obtain copies of these studies if they did. Finally, we conducted searches of the websites of foundations, research organizations, and governmental agencies; in all, more than 70 were searched (see Table 1). Based on our sense that research on this topic often does not appear in peer-reviewed outlets, we believed that this was a critical step in locating relevant studies (see the section "Publication Bias" for a description of the problems associated with relying solely on peer-reviewed work, as well as efforts that we made to assess the extent to which publication bias might be affecting our results). After obtaining electronic or physical copies of the studies that passed through our initial screen, we utilized a secondary screening process to determine if studies actually met our inclusion criteria. As such, studies that looked promising based on a reading of their titles and abstracts were evaluated for inclusion using the full study text, as is often recommended (see Cooper, 1998).

Coding Studies

After the second screening, all studies that seemed to meet our inclusion criteria were fully coded. We coded background study characteristics (e.g., authors, year of publication), characteristics of the intervention (e.g., the type of transition addressed, the duration of the intervention), characteristics of the sample (e.g., age, source of risk status), and outcomes (e.g., construct measured, effect size). Again, all studies were coded by two authors working independently, and differences among coders were resolved through discussion.

Quality assessment of included studies. We attempted to assess, as part of the coding, the likely internal, external, construct, and statistical validity of the inferences arising from all studies. We approached this aspect of the review using the framework provided by Valentine and Cooper (2008). This approach attempts to overcome many of the shortcomings of existing quality

TABLE 1

Websites Searched for Possible Relevant Studies

Organization	Internet address
Academic Pathways to Access Student Success	www.apass.uiuc.edu
Academy for Educational Development	www.aed.org
Achieve, Inc.	www.achieve.org
Association for Career and Technical Education Research	www.agri.wsu.edu/acter/
American Association for Community Colleges	www.aacc.nche.edu
American Institutes for Research	www.air.org
American Youth Policy Forum	www.aypf.org
Annie E. Casey Foundation	www.aecf.org
Association for Supervision and Curriculum Development	www.ascd.org
Association for Career and Technical Education	www.acteonline.org
Association of American Colleges and Universities	www.aacu.org
Berea College	www.berea.edu
Broad Foundation	www.broadfoundation.org
Carnegie Foundation	www.carnegiefoundation.org
Career Education Corporation	www.careered.com
The Center for Community College Policy	www.communitycollegepolicy.org
The Center for Educational Policy Research	www.s4s.org
Center for Occupational Research and Development	www.cord.org
The Center for Research on Developmental Education and Urban Literacy	www.cehd.umn.edu/crdeul/about.html
Center on Education and Work	www.cew.wisc.edu/
College of the Ozarks	www.cofo.edu
Community College Resource Center	www.ccrc.tc.columbia.edu/
Council for Exceptional Children	www.cec.sped.org//AM/Template.cfm?section=Home
ED Publications (U.S. Department of Education)	www.ed.gov/about/pubs/intro/index.html?src=gu
Education Commission of the States	www.ecs.org
Educational Policy Institute	www.educationalpolicy.org
The Education Trust	www2.edtrust.org/edtrust/
Ford Foundation	www.fordfound.org
Bill and Melinda Gates Foundation	www.gatesfoundation.org/Pages/home.aspx
The Institute on Education and the Economy	www.tc.columbia.edu/centers/iee/
Jobs for the Future	www.jff.org
W.K. Kellogg Foundation	www.wkkf.org
Latin America Research and Service Agency	www.larasa.org
League for Innovation in the Community College	www.league.org/index.cfm
League of United Latin American Citizens	www.lulac.org
Lilly Foundation	www.lilly.com/products/access/foundation.html
Lumina Foundation	www.luminafoundation.org
Ronald E. McNair Postbaccalaureate Achievement Program (TRIO)	www.ed.gov/programs/triomcnair/index.html
MDRC (Manpower Demonstration Research Corporation)	www.mdrc.org
Mathematica Policy Research, Inc.	www.mathematica-mpr.com/
MELMAC Education Foundation	www.melmacfoundation.org
Charles Stewart Mott Foundation	www.mott.org
MPR Associates	www.mprinc.com
National Association for the Advancement of Colored People	www.naacp.org/home/index.htm
National Career Pathways Network	www.cord.org/ncpn-index.cfm
National Center for the First-Year Experience and Students in Transition	www.sc.edu/fye/
National Center for Public Policy and Higher Education	www.highereducation.org/index.shtml
National Council for Workforce Education	www.ncwe.org
National Governors Association	www.nga.org
National Research Center for Career and Technical Education	www.nrccte.org
National Center for Research in Vocational Education	vocserve.berkeley.edu
Nellie Mae Education Foundation	www.nmefdn.org
National Science Foundation	www.nsf.gov
Office of Community College Research and Leadership	occrl.ed.uiuc.edu/
Office of Special Education Programs	www.ed.gov/about/offices/list/osers/osep/index.html
Office of Vocational and Adult Education	www.ed.gov/about/offices/list/ovae/index.html
Frederick D. Patterson Institute	http://www.patterson-uncf.org/index.htm
Pew Charitable Trust	www.pewtrusts.org
Postsecondary Education Opportunity	www.postsecondary.org
Rockefeller Foundation	www.rockfound.org
United Negro College Fund	www.uncf.org
Upjohn Institute for Employment Research	www.upjohninst.org
Washington Center for Improving the Quality of Undergrad Education	www.evergreen.edu/washcenter/project.asp?pid=73
Washington State Board for Community and Tech Colleges	www.sbctc.ctc.edu
WestEd	www.wested.org/cs/we/print/docs/we/home.htm
Women Employed	www.womenemployed.org
Workforce Strategy	www.workforcestrategy.org/

scales. Among these are the reliance on single scores to represent the multidimensional construct of study quality and the use of items that require a great deal of inference on the part of the study coders. Specifically, we coded questions that addressed the internal validity (e.g., how participants were assigned to conditions, overall attrition rate), external validity (e.g., degree to which the sample appeared to be representative of the target population), construct validity (e.g., reliability of scores), and statistical validity (e.g., the extent to which data met assumptions underlying the general linear model). As will be shown later, however, due to poor reporting in the primary studies, we were unable to assess studies on most quality dimensions.

Data Analysis

To analyze data, we conducted a meta-analysis treating each independent sample within the studies as the unit of analysis (most studies provided only one independent sample). We weighted effect sizes by the amount of information they provided about the population mean (i.e., we used the typical inverse variance weight; Shadish & Haddock, 2009) and calculated 95% confidence intervals for all effects. In addition, this data analysis strategy required us to make several decisions about how we would address certain complexities in the data. Each of these is described below.

Effect size metric. We employed effect sizes that reflect mean differences between groups for continuous outcomes. In addition, because we anticipated that the authors of most studies would choose to operationalize their constructs in different ways, we standardized these effect sizes to provide an interpretable comparison across studies. The standardized mean difference effect size is computed as

$$d = \frac{Y_T - Y_C}{s_p}, \qquad (1)$$

where $\bar{Y}_T$ is the treatment group mean, $\bar{Y}_C$ is the comparison group mean, and s_p is the pooled standard deviation. This formulation of the standardized mean difference effect size has a known bias in small samples, so we applied the usual correction for this (Hedges, 1981). Standardized mean difference effect sizes were computed so that values greater than zero indicated positive program effects (e.g., better grades for the intervention group relative to the comparison group).

For dichotomous outcomes (e.g., persisted vs. did not persist), we employed the odds ratio, which is defined as

$$OR = \frac{\frac{a}{b}}{\frac{c}{d}} = ad \quad bc, \qquad (2)$$

where a is defined as the number of successes in the intervention group, b is the number of nonsuccesses in the intervention group, c is the number of successes in the comparison group, and d is the number of nonsuccesses in the comparison group. Meta-analysis was performed on logged odds ratios, because this metric has better statistical properties (Lipsey & Wilson, 2001). We then transformed the logged odds ratios to odds ratios for presentation. The odds ratios were set up so that values greater than one indicated positive program effects (e.g., greater persistence rates among intervention group members relative to comparison group members).

Dependencies among effect sizes. Independence of observations is a fundamental assumption underlying the general linear model, but in a meta-analysis, effect sizes (observations) can be dependent for several different reasons. For example, researchers may assess an outcome at the end of a program and at a subsequent follow-up. Or, researchers might collect multiple measures of the same construct (e.g., scores on a standardized test and class grades as a measure of academic achievement). Because these effects are based on the same sample, they are not independent, and treating them as such has undesirable effects. Chief among these are that studies will not be weighted properly in a meta-analysis, and statistics that estimate variability (or are based on estimates of variability) will be biased.

To address the problem of dependence, researchers conducting meta-analyses have choices. One is to randomly select one measure to represent the study. This has the virtue of dealing with the dependence problem but the

drawback of discarding information. Another strategy is to select the outcome that maximizes similarity with other studies. Often this strategy is better than randomly selecting one effect size, but it still results in a loss of information. A final strategy is often referred to as the shifting unit of analysis approach (Cooper, 1998); this is the approach we adopted for this article. The shifting unit of analysis approach involves averaging effects when appropriate (e.g., averaging posttest and follow-up effects when testing the overall effect size for the intervention), then splitting the effects when testing the dimension on which they differ. In this example, when asking whether program effects appear to persist over time, a study with both a posttest effect size and a follow-up effect size would contribute to both levels of that moderator.

Error model. Finally, researchers need to consider whether to employ a fixed effect or a random effects analytic model. Using the fixed effect model, study effects can be thought of as being randomly sampled from a single population of studies, and therefore any differences in effect sizes across studies are treated as solely due to random sampling and identifiable covariates. Using the random effects model, reviewers assume that studies do not in fact share a single population value, and any differences in between-study effect sizes are due to random sampling error, any identifiable covariates, and other random factors that cannot be identified.

The choice between fixed and random effects models can be an important one, because the confidence intervals arising from a random effects analysis will never be smaller and are often larger than their fixed effect counterparts; this has implications for both the statistical significance tests and interpreting the likely range of an intervention's effect. In practice, choosing between the models is done empirically or conceptually. Empirically, reviewers often allow a statistical test of homogeneity to dictate their choice. The formula for this test is

$$Q = \frac{w_i(ES_i - \overline{ES})^2}{k-1}, \qquad (3)$$

where ES_i is each individual effect size, w_i is the inverse variance weight for effect size i, and $\overline{ES}$ is the weighted average effect size. Formally, the homogeneity test is a test of the between-studies variance component, and a significant value of Q indicates that the variation among studies is significantly different from zero, often leading researchers to employ a random effects model. Generally, however, the choice between models is best made conceptually (Hedges & Vevea, 1998; Valentine et al., 2010). Researchers might, for example, consider the diversity of research designs, programs, outcomes, and samples in their review and use a fixed effect model if these seem to be very similar. Or, they may have an explicit interest in the extent to which different studies yield different answers. In addition, researchers could consider their desired universe of generalization as a basis for choosing between the models. The fixed effect model allows for generalization to studies that are highly like the ones in the review, whereas the random effect model allows for inferences that are not so tightly conditioned on the observed studies (hence, the inferences are more broadly generalizable). In this review, the programs meta-analyzed were not highly similar; in addition, we were interested in the extent to which the studies seemed to yield different effects, and we also wanted to take advantage of the broader range of generalization. As such, we adopted the random effects data analytic model.

Results

Literature Search

Because we cast a very wide net with our electronic search, we identified more than 8,000 possible citations. The vast majority of these were not relevant to the purposes of our study (e.g., were opinion pieces, literature reviews, or simple descriptions of transition programs), and only 109 were selected for further evaluation. Of these, 33 were selected for full coding as possibly meeting inclusion criteria; on further examination, some of these turned out not to meet our inclusion criteria. The search of the literature reviews yielded an additional 149 potential studies, and the website search yielded an additional 30. As with the electronic literature search, most of these

could not be included in this review. From all sources, a total of 19 unique studies meeting our criteria were identified; each of these studied interventions aimed at helping at-risk students remain in college. Both community colleges and 4-year institutions were represented, but the latter was by far the largest category. In the analyses below, we combined these different institutions because the approaches used to help students were similar. For 18 of these studies, we were able to compute an effect size that described the impact the intervention had on program participants. These studies are described in Table 2 (for achievement outcomes) and/or Table 3 (for retention outcomes). Each table provides a brief description of the main elements of the intervention, the target population and setting, the duration of the intervention, the comparison group, the specific outcome, the timing of the measurement of the outcome, and the effect size.

Description of Interventions

As anticipated, the interventions included in this review varied in interesting ways. Because the interventions were designed to help college students stay in school (i.e., are college transition programs), they all included students who were either at increased risk for college failure (e.g., were identified as high-risk admits) or were on academic probation. Due to their common purpose, the studies were similar in that they included students with a variety of background characteristics (e.g., ethnicity, income) and included both men and women. However, the specific approaches taken in these studies varied quite a bit. These ranged from relatively comprehensive interventions (e.g., a seminar designed to facilitate college adjustment, coupled with limitations on the number of credit hours students could enroll in, smaller classes, and tutoring; Hecker, 1995) to those that were much smaller in scale, such as adding a journaling component to an English composition class (Cohen Goodman, 1998). Most interventions fell between these two poles, with a freshman orientation/adjustment seminar being the strategy most often adopted (either alone or in conjunction with other activities such as tutoring).

Description of Research Designs and Study Implementation

Unfortunately, only three of the 19 studies included in this review used a random mechanism to assign students to groups (these were Cohen Goodman, 1998; Dees, 1991; Scrivener et al., 2008). As such, selection (i.e., differences in participants who receive the intervention relative to those who do not) is a pervasive threat to the validity of these studies. Notably, one additional study (Moss & Yeaton, 2006) employed a regression-discontinuity design to study the effects of a developmental (or remedial) English program for college students. This study was not included in the meta-analysis due to the difficulties of converting the statistical results of the regression-discontinuity design to a metric compatible with the rest of the studies in this review, but it will be discussed below.

Potential selection effects were addressed to varying degrees in the studies included in the meta-analysis. Most of these studies compared the students receiving the intervention to students who appeared to be comparable but did not receive the intervention. The best of these studies attempted to adjust for baseline differences between groups or allowed us to accomplish the same goal. For example, some studies computed means that were adjusted for scores on college entrance exams, whereas others allowed us to compute a pretest effect size that was subtracted from the reported posttest effect size. However, some studies compared program students to students who were clearly not comparable (e.g., Cone, 1991). We explore the consequences of these design choices below.

A pervasive threat to the statistical conclusions reached in the individual studies, and relevant to the confidence intervals we generated through meta-analyses of them, is that many of the included studies likely violated the statistical assumption of independence of observations. For example, many studies involved comparing students in some sections of a course to students in other sections of a course. Because students were not randomly assigned to sections, shared characteristics of students choosing the same section may increase the similarity of observations within sections relative to the similarity of observations across sections. Furthermore, students in the

TABLE 2
Studies Measuring Academic Achievement Outcomes

First author (year)	Intervention description	Target population/ setting	Duration of intervention	Comparison group	Outcome	Outcome assessment timing	Effect size[a]
Abadie (1999)	Administrative limitations on extracurricular activities, smaller class sizes, required body of general education courses	Incoming 4-year college students not meeting regular admit criteria	One academic year	Students admitted via usual admission process	GPA	After first semester of intervention After second semester of intervention	–0.13 –0.56
Alderman (1998)	One credit college orientation class, tutoring, remedial coursework	Community college students identified as needing remedial instruction	One semester	Historical controls who met inclusion criteria (i.e., students before program existed)	GPA	Semester following the end of the intervention	+0.18
Clark (1993)	Remedial coursework, small classes, academic and career advising	Incoming 4-year students scoring in the lowest quartile of a placement test	One academic year	Historical controls who met inclusion criteria (i.e., students before program existed)	GPA	Three and a half years after the end of the intervention	+0.93
Cohen Goodman (1998)	Added a journal writing component to an English composition class	Students in a 4-year university scoring low on a placement test	One semester	Randomly assigned group of students not assigned to instructors in the intervention condition	Reading comprehension	Immediately after intervention	+0.07
Cone (1991)	Study skills and adjustment course	Students in a 4-year university with first-semester GPA < 2.0	One semester	Unclear, but comparison students do not appear to have met inclusion criteria for intervention	GPA	Immediately after intervention	–0.61
Cox (2002)	Study skills curriculum integrated into usual math instruction	Students in a community college scoring below cutoff on a placement test	One semester	Historical controls who met inclusion criteria (i.e., students before program existed)	Grade in one math class	Immediately after the intervention	+0.32
Dees (1991)	Cooperative learning in a remedial math class	Students in a 4-year university needing remediation	One semester	Randomly assigned group of students in a traditional lab	Various math tests	Immediately after the intervention	+0.37

Esterbrook (2006)	Behavior modification	Students in a community college subjectively deemed to be "at-risk"	Unclear, but appears to be one semester	Assigned to receive traditional remedial instruction	GPA	Immediately after the intervention	–0.30
Fry (2007)	Course aimed at fostering time management and problem-solving skills, as well as increased awareness of university resources	Conditionally admitted students in a 4-year university	One semester	Conditionally admitted students not taking the seminar Unconditionally admitted students not taking the seminar	GPA	Unclear, but appears to be for the semester during which the seminar took place	+0.21 –0.54
Hecker (1995)	Administrative limitations on maximum credit hours, courses available, class sizes; seminar to teach academic skills	Students conditionally admitted to a 4-year university through an alternate process	Probably 1 academic year	Regularly admitted students deemed to be "high risk" Regular admits	GPA	Immediately following the end of the intervention (i.e., Fall semester of the Sophomore year)	+0.15 –0.58
Loiacano (2000)	Specific curriculum added to an existing freshman orientation course	First-year students in a 4-year university with a history of academic struggles	One academic semester	Similar students not receiving the added curriculum	Cognitive development	Immediately after the end of the intervention	+0.11
McGee (2004)	Statewide program providing financial, academic, and social assistance	Disadvantaged students attending community college	Unclear, but presumably 2 years	Students not participating in the program	GPA	Unclear	–0.22
McGregor (2001)	Added component to an existing college preparation course	Entering freshmen not meeting usual admission criteria to a 4-year university	Five weeks	Similar students not receiving the added component	Vocabulary Critical thinking	Immediately after the end of the intervention Immediately after the end of the intervention	+0.16 +0.16
Milligan (2007)	Study skills seminar	Students on academic probation in a 4-year university	Eight weeks	Similar students who chose not to participate in the seminar	GPA	Immediately after the end of the intervention	+0.22

(continued)

TABLE 2 (continued)

Salinitri (2005)	Mentoring	Students with entrance scores near the institution's lower limit	Unclear, but probably one semester	Similar students not chosen to receive mentoring	GPA	Unclear, but probably refers to the semester immediately following the intervention	+0.60
Sanders (2000)	Peer tutoring	Academically underprepared freshmen in a 4-year institution	One academic year	Unclear, but seems to be similar students who did not receive the intervention	GPA	Immediately after the end of the intervention	+0.61
Scrivener (2008)	Learning communities of about 25 students; each community took a set of three courses together; curricula across the courses were linked. Tutoring was also offered	Freshmen at a community college (sample in the meta-analysis had failed both English placement tests given by the college)	One semester	Students randomly assigned to receive college's usual menu of courses and support	Whether students had passed an English course[b]	First full semester after the end of the intervention	+0.22
Stovall (1999)[c]	Student success course focusing on transitioning to college, career development, and life management	Students in a community college	One semester	Students who scored below college level on two placement tests (reading and English)	GPA	Immediately after the end of the interventiond	+0.21

a. All effect sizes for academic achievement are expressed as standardized mean differences. A standardized mean difference > 0 indicates that the students receiving the intervention performed better than students in the comparison condition.

b. In Scrivener et al. (2008), this outcome was expressed in terms of percentages of students who had passed an English course versus those who had not. We computed a logged odds ratio for this outcome, then transformed that logged odds ratio to a standardized mean difference effect size.

c. Stovall (1999) did not analyze students who took but did not pass the student success course. This choice may positively bias the effect size estimate somewhat.

d. Stovall (1999) also measured GPA at the end of the second term, 2nd academic year, and 3rd academic year. She found no differences between program participants and nonparticipants and did not separately compare at-risk program participants to at-risk nonparticipants. For meta-analysis, we conservatively imputed 0 for these effects.

TABLE 3
Studies Measuring Persistence Outcomes

First author (year)	Intervention description	Target population/setting	Duration of intervention	Comparison group	Outcome	Outcome assessment timing	Effect size[a]
Abadie (1999)	Administrative limitations on extracurricular activities, smaller class sizes, required body of general education courses	Incoming 4-year college students not meeting regular admit criteria	One academic year	Regular admits	Retention	First semester after intervention end Second semester after intervention end	0.36 0.42
Alderman (1998)	One credit college orientation class, tutoring, remedial coursework	Community college students identified as needing remedial instruction	One semester	Historical controls who met inclusion criteria (i.e., students before program existed)	Retention	First semester after intervention end	1.32
Clark (1993)	Remedial coursework, small classes, academic and career advising	Incoming 4-year students scoring in the lowest quartile of a placement test	One academic year	Historical controls who met inclusion criteria (i.e., students before program existed) Regular admits	Retention	Three and a half years after the end of the intervention	1.21 0.86
Cone (1991)	Study skills and adjustment course	Students in a 4-year university with first-semester GPA < 2.0	One semester	Historical controls who met inclusion criteria (i.e., students before program existed)	Retention	Semester following the intervention	14.48
Fry (2007)	Course aimed at fostering time management and problem-solving skills, as well as increased awareness of university resources	Conditionally admitted students in a 4-year university	One semester	Conditionally admitted students not taking the seminar Unconditionally admitted students not taking the seminar	Retention	Semester following the intervention Two semesters following the intervention Semester following the intervention Two semesters following the intervention	1.12 0.88 0.76 0.64

(continued)

TABLE 3 (continued)

Hecker (1995)	Administrative limitations on maximum credit hours, courses available, class sizes; seminar to teach academic skills	Students conditionally admitted to a 4-year university through an alternate process	Probably 1 academic year	Regularly admitted students deemed to be "high risk" Regular admits	Retention Retention	Immediately following the end of the intervention (i.e., Fall semester of the sophomore year)	0.77 0.49
House (1991)	Tutoring	Academically underprepared freshmen admitted through a special process to a 4-year university	Probably 1 academic year	Students eligible to receive tutoring but who did not	Retention	Appears to be the next semester after the end of the intervention (i.e., Fall semester of the sophomore year)	1.52
Milligan (2007)	Study skills seminar	Students on academic probation in a 4-year university	Eight weeks	Similar students who chose not to participate in the seminar	Retention	Immediately after the end of the intervention	1.15
Salinitri (2005)	Mentoring	Students with entrance scores near the institution's lower limit	Unclear, but probably one semester	Similar students not chosen to receive mentoring	Retention	Unclear, but probably refers to the semester immediately following the intervention	14.60
Sanders (2000)	Peer tutoring	Academically underprepared freshmen in a 4-year institution	One academic year	Unclear, but seems to be similar students who did not receive the intervention	Retention	Immediately after the end of the intervention	0.83
Stovall (1999)[b]	Student success course focusing on transitioning to college, career development, and life management	Students in a community college	One semester	Students who scored below college level on two placement tests (reading and English)	Retention	Immediately after the end of the intervention End of the second academic year End of the third academic year	23.69 1.94 1.39

a. All retention effect sizes are expressed as odds ratios. An odds ratio > 1 indicates that the intervention was associated with increased retention.
b. Stovall (1999) did not analyze students who took but did not pass the student success course. This choice may positively bias the effect size estimate somewhat.

same section share other influences, such as the instructor, that likely also tend to increase their relative similarity. Most important for the purposes of this review, violation of the assumption of independence can lead to standard errors that are spuriously small—and hence confidence intervals that are too narrow and statistical tests that are too likely to reject the null hypothesis. Unfortunately, few of the studies in this review gave us enough information about the nature and extent of potential data clustering, and none allowed us to estimate a likely value that could be used to arrive at potentially better standard errors.

Almost without exception, the studies gave little indication about other potential threats to their validity. For example, most studies did not provide information on attrition (either overall or differential), data exclusions (such as systematically missing data), or intervention fidelity. The absence of good reporting about these issues means that we do not know how serious these threats are to the validity of the conclusions we draw below.

Outcomes Measured

Most studies measured academic achievement (usually via grade point average [GPA]) or persistence (as measured by reenrollment). The majority of studies measured these outcomes either immediately after the program (e.g., for a Fall 2000 course, reenrollment for the Spring 2001 semester), or one semester later. Only two studies (Clark & Halpern, 1993; Stovall, 1999) can be considered to have measured outcomes over the long term (i.e., after more than 1 year).

Data Dependencies

Abadie (1999) reported outcomes for two cohorts of students. We collapsed these into one group for analysis purposes. Abadie (1999) and Stovall (1999) also reported outcomes for multiple points in time (Abadie, immediately following the end of the intervention and one semester later; Stovall, at the end of the first and second semesters, the end of the 2nd year, and the end of the 3rd year). We also collapsed these for analysis. We would have employed Cooper's (1998) shifting unit approach here if more studies had measured outcomes at similar follow-up periods, but the lack of this feature across studies made such an analysis impossible. McGregor (2001) administered two measures related to academic achievement, and these were averaged for analysis. Finally, two studies (Fry, 2007; Hecker; 1995) used two comparison groups (one comparison group that was made up of regularly admitted students and a second group that was made up of students who were more like the students receiving the intervention). Here, we did employ Cooper's shifting unit of analysis approach and averaged these groups for the overall analysis, but we allowed them to contribute to both levels of analysis when we examined the nature of the comparison group and its effects on effect size estimation.

Data Analysis

Academic achievement. Eighteen studies measured the impact of an intervention on academic achievement (most often GPA; see Table 2). The random effects estimate was positive, indicating that program participants fared better on achievement related outcomes, but not statistically significant, $d = .08 \pm .17$, $p = .30$. The distribution of effect sizes was heterogeneous, $Q(17) = 68.7$, $I^2 = 75\%$, $p < .001$, suggesting that the study effects varied more than would be expected given chance alone.

Persistence. Eleven studies measured the impact of an intervention on student persistence (see Table 3). The random effects estimate for this outcome was positive (indicating that program participants were more likely to reenroll) but not statistically significant, with an odds ratio of 1.46 (the 95% confidence interval ranged from .85 to 2.51), $p = .17$. Once again, the distribution of effect sizes was heterogeneous, $Q(10) = 84.7$, $I^2 = 88\%$, $p < .001$.

Publication bias. Publication bias—the tendency for studies lacking statistically significant outcomes to go unpublished, and therefore for reviews relying only on peer-reviewed work to overestimate the relations being studied—is a concern in every review, even those that, like ours, include a vigorous search for unpublished literature. For the academic achievement outcomes, we conducted a statistical analysis to help assess whether our set of

studies appeared to be affected by publication bias. It should be noted that there are no very good solutions to the problems posed by publication bias, and current statistical approaches are at best informed guesses about the nature and severity of potential publication bias. We used the trim and fill approach, which is based on the assumption that the observed studies (i.e., those in the meta-analysis) are a random sample from a normally distributed population of studies (see Borenstein, Hedges, Higgins, & Rothstein, 2009, for a discussion of this technique).

Our analysis suggests that, in fact, some degree of publication bias might exist in our set of achievement outcomes. Specifically, the trim and fill analysis identified that an additional two studies would need to be added to the distribution of effect sizes in the achievement meta-analysis for that distribution to be essentially normal. However, adding these studies does not substantively alter the interpretation of the overall meta-analysis on achievement outcomes. Specifically, the new estimated effect size is still small and not statistically significant, $d = .01$, $p = .98$, and is neither statistically significantly nor practically significantly different from our overall estimate. As such, the hints of publication bias that exist do not seem to suggest a great deal of concern about the integrity of our meta-analytic data set.

Comparison quality as a moderator of study effects. For both grades and persistence, the distributions of effect sizes were heterogeneous. Heterogeneity is one justification for the search for moderating influences, as these might explain some or potentially all of the "excess" observed heterogeneity. We noted that studies varied in terms of the quality of the comparison group against which the relative effects of the intervention were judged. We therefore investigated whether the quality of the comparison group moderated the effect sizes we observed. This was very much the case. The five comparisons of the academic achievement of program students to clearly noncomparable students yielded a strong, negative, statistically significant program effect ($d = -.45 \pm .17$, $p < .001$), whereas the 15 comparisons of program students to relatively more comparable students yielded a statistically significant positive effect ($d = .25 \pm .11$, $p < .001$). A fixed effect moderator test for the difference between these two groups of effect sizes was statistically significant, $Q(1) = 68.1$, $p < .001$. Notably, effects within each level of comparability appeared to be relatively homogeneous (both ps $> .16$), suggesting that the programs within each category were associated with relatively similar effect sizes.

For persistence outcomes, the nature of the comparison group again moderated the observed effect sizes. The four studies that compared the persistence of program students to clearly noncomparable students yielded negative effects (the weighted average odds ratio was 0.69, with a 95% confidence interval ranging from 0.34 to 1.40), whereas the 10 comparisons of program students to relatively more comparable students yielded positive effects (the weighted average odds ratio was 1.21, with a 95% confidence interval ranging from 1.001 to 1.46). The fixed effect moderator test for the difference between these two groups of effect sizes was statistically significant, $Q(1) = 24.4$, $p < .001$. Effects within each level of comparability again appeared to be relatively homogeneous (both ps $> .29$).

"Best practice" studies relative to other studies. A final comparison of interest involves those studies that, compared to the other studies in our review, can be considered "best practice." That is, these studies used both a relatively intensive intervention and a relatively better comparison group. As expected, the nine comparisons with these characteristics yielded positive and statistically significant effect sizes for academic achievement ($d = .29 \pm .15$, $p < .001$), whereas the 11 comparisons that lacked either an intensive intervention, a relatively good comparison, or both yielded negative effect sizes in the random effects model ($d = -.17 \pm .21$, $p = .12$). The fixed effect moderator test for the difference between these two groups of effect sizes was statistically significant, $Q(1) = 44.5$, $p < .001$. Studies in the "best practice" category appeared to be relatively homogeneous ($p = .24$, $I^2 = 23\%$), whereas studies in the lower quality category were still heterogeneous ($p < .001$, $I^2 = 70\%$). There was not enough variation on

the quality dimension to do a similar analysis for studies that measured persistence, although an examination of Table 3 suggests that the pattern is similar.

Interpreting the program effects. If we assume that the best estimate of the effects of transition programs on student achievement comes from the best practice studies, this suggests that these programs have a population effect of about $\delta = .29$ on student grades. To put this in context, assume that students are expected to have a GPA of 2.0 if they do not receive the intervention. A population effect of $\delta = .29$ implies that students receiving the intervention should be about three-tenths of a standard deviation higher in GPA than those not receiving the intervention. A typical standard deviation for GPA in the lower portion of the distribution is about .75, so the typical student receiving the intervention should have a GPA of about 2.22 [i.e., $2.0 + (.75 \times .29)$]. For every five program students taking 12 credit hours, this would be approximately equivalent to four of them earning three Cs and one B, with the fifth earning four Cs, whereas the five students in the comparison condition would earn all C grades.

One problem with the interpretation of the program effects for grades is that some interventions required students to attend a seminar or course, and it was not always clear if or how these seminars were graded. It was also not always clear how many credit hours students would have enrolled for, if these were formal courses. If the seminars were graded and this grade was included in the computation of the GPA, and if the seminar counted for 3 credit hours, then the observed effect on GPA could largely or entirely be due to the influence of the program course on grades. Clearly this is an issue that merits specific attention in future studies.

To interpret the persistence outcomes, assume that the "true" intervention effect is given by the odds ratio for the studies that used a relatively better comparison group (i.e., the odds ratio for persistence is 1.62). If we assume—optimistically—that about 50% of students would reenroll the next semester in the absence of the intervention, then the odds ratio suggests that for about every 10 students who receive the intervention, one additional student would persist the next semester. Of course, due to the fact that the studies included in this review tended not to measure outcomes beyond two semesters after the intervention, we know very little about how program effects behave over time.

A study employing the regression-discontinuity design. Moss and Yeaton (2006) used regression-discontinuity to study the effectiveness of a developmental English program in a large community college. The regression-discontinuity design involves assigning a cutoff point, below (or above) which participants receive the intervention. For example, a pretest might be administered, and all potential participants falling below a certain threshold score might be given the intervention. This design generally has very strong inferential properties, as, like studies using random assignment, the selection mechanism is entirely known.

In the Moss and Yeaton (2006) study, the college administered a placement exam in English. Students falling below a certain threshold score were required to take a developmental English course before they could take college-level English. The outcome variable was the grade that the students earned in their college-level English course. Results suggested that the developmental English course was associated with better grades in the college-level course, although this effect appears to have been concentrated in the students who scored the lowest on the placement test. In other words, students in the developmental course who performed relatively well on the placement exam (i.e., those right below the cut-point requiring the course) received grades similar to students who fell just on the other side of the cut-point. However, students who scored very low on the placement exam received grades in their college-level course that were similar to their peers who scored higher on the placement exam. Follow-up analyses suggested that neither differential maturation nor differential attrition, both potential rival hypotheses in this particular study, appear to have influenced study results.

Discussion

The results of this review suggest that there is reason to be optimistic about the effects of programs designed to help at-risk college students stay in school, at least in the short term. That is, as a group, interventions that were relatively more intensive and methodologically sounder had outcomes that suggested small, short-term benefits are associated with participation in such programs. However, the studies we reviewed provide an unfortunately weak basis for public policy, because their designs tend not to be strong. Researchers studying college student persistence often use extant databases, such as those maintained by institutions for data management purposes. While such secondary data analyses can be useful, particularly in the early stages of the research process, the available variables are limited and often constrain theoretical development and testing. This is particularly true in the context of understanding a student's experience off campus, their attitudes and perceptions of the college experience, and their educational goals. Similarly, the dearth of rigorously designed studies that examine postsecondary transition programs compromises the potential for theoretical advances in understanding what works for at-risk students.

Furthermore, the studies included in this systematic review tended to lack reporting on details that would allow us to assess the conditions under which, and characteristics of students for whom, the interventions might be effective. As an example, due to poor reporting, we were unable to critically examine the quality of the included studies in a rigorous manner. More specifically, few of the included studies discussed implementation fidelity in much depth, and as such we have little information about the degree to which observed effects might be attenuated due to low fidelity. Also, most studies employed an evaluation design in which students were allowed to choose whether they received the intervention or a comparison condition (e.g., a voluntary course or an analysis of a database that tracked student experiences). Although the researchers often (but not always) took steps to attempt to make the intervention and comparison groups more comparable, these designs still carry with them an added element of ambiguity. This concern partly exists because we can never know how well our attempts to make groups more comparable have worked. Future studies using this type of design and analysis strategy should attend more explicitly to the concerns raised by nonexperimental designs. In addition, more detailed reporting of quality characteristics is critical to understanding the reported effects of studies of these programs. For example, one would think that dropping out of the study would be a fundamental concern in studies of interventions designed to increase student persistence, but study attrition was often not reported and only rarely investigated as a threat to validity in the studies in this review.

Even though the data seem to suggest that, among our stronger studies, there is evidence that the comprehensive interventions might positively affect short-term grades and persistence, we have little information about which elements in the comprehensive interventions might be relatively more effective. Funding agencies could help improve the state of the research by funding more rigorous research but even more importantly by ensuring that future evaluations of retention programs describe in detail the nature of the intervention. In addition, it will be helpful for researchers to collect and report details about resource utilization and other factors related to implementation fidelity. For example, several interventions included in our review had a tutoring component, but no studies provided detailed information about the training of tutors or the number of tutoring sessions attended. Information of this kind is critical for understanding what happened in the study (so that the program can be replicated in other locations), but also for potentially explaining different results across studies (e.g., why one study determined that tutoring was effective and another did not).

Finally, most studies contained little information that would help individuals make decisions about how best to support students in particular areas, such as those in career and technical education or those in community college settings, and most provided virtually no information about program costs. Taken together, these concerns suggest potentially serious gaps in our understanding of the effectiveness of specific program elements to support educational transitions for at-risk students. Our hope is that this review spurs

rigorous and theoretically rich studies of interventions that aim to support at-risk students in their efforts to graduate from college and further the public and private benefits that accrue from higher levels of educational attainment.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abadie, M. M. (1999). Comparison of students enrolled in an alternative academic program with regularly enrolled students in a Research I university. *Dissertation Abstracts International*, *60*(3), 666A.

*Alderman, L. V. (1998). Student support services and their impact on persistence of first-year students at a rural community college. *Dissertation Abstracts International*, *58*(11), 4201A.

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, *25*, 297–308.

Barrow, L., & Rouse, C. E. (2005). Does college still pay? *Economists' Voice*, *4*, 1–9.

Barton, P. E. (2008, January/February). How many college graduates does the U.S. labor force really need? *Change*, *40*, 16–21.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, *12*, 155–187.

Bean, J. P. (1983). The application of a model of turnover in work organizations to the student attrition process. *Review of Higher Education*, *6*, 129–148.

Bean, J. P., & Metzner, B. M. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, *55*, 485–540.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis with special reference to education* (3rd ed.). Chicago, IL: University of Chicago Press.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York, NY: Greenwood.

Bourdieu, P., & Passeron, J. C. (1977). *Reproduction in education, society, and culture*. Beverly Hills, CA: Sage.

Braxton, J. M. (Ed.). (2000). *Reworking the departure puzzle*. Nashville, TN: Vanderbilt University Press.

Braxton, J. M., & Mundy, M. E. (2001–2002). Powerful institutional levers to reduce college student departure. *Journal of College Student Retention*, *3*, 91–118.

Braxton, J. M., Hirschy, A. S., & McClendon, S. A. (2004). *Understanding and reducing college student departure.* San Francisco: Jossey Bass.

Braxton, J. M., Sullivan, A. S., & Johnson, R. M., Jr. (1997). Appraising Tinto's theory of college student departure. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 12, pp. 107–164). New York, NY: Agathon.

Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equation modeling test of an integrated model of student retention. *Journal of Higher Education*, *64*, 123–139.

Carnevale, A. P., & Derochers, D. M. (2003). *Standards for what? The economic roots of K-16 reform*. Princeton, NJ: Educational Testing Service.

*Clark, J. M., & Halpern, D. F. (1993). The million dollar question: Can an intensive learning experience help lowest-quartile students succeed in college? *Journal of Instructional Psychology*, *20*, 29–39.

Clark, J., Hernandez, A., & Burkey, C. L. (2007). *Children of immigrants and education in Indiana: Reports from the frontlines*. Indianapolis, IN: Sagamore Institute for Policy Research.

*Cohen Goodman, M. (1998). The effect of journal writing on the reading comprehension, study habits and attitude of underprepared college students. *Dissertation Abstracts International*, *59*(3), 722A.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, 95–120.

*Cone, A. L. (1991). Sophomore academic retention associated with a freshman study skills and college adjustment course. *Psychological Reports*, *69*, 312–314.

Cook, T. D., Shadish, W. R., & Wong, V. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management*, *27*, 724–750.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.

*Cox, F. W. (2002). The relationship of study skills and mathematics anxiety to success in mathematics among community college students. *Dissertation Abstracts International*, *62*(10), 3325A.

*Dees, R. L. (1991). The role of cooperative learning in increasing problem-solving ability in a college remedial course. *Journal for Research in Mathematics Education*, *22*, 409–421.

Eaton, S. B., & Bean, J. P. (1995). An approach/avoidance behavioral model of college student attrition. *Research in Higher Education*, *36*, 617–645.

*Esterbrook, R. L. (2006). Introducing Russian neurolinguistic programming behavior modification techniques to enhance learning and coping skills for high-risk students in community colleges: An initial investigation. *Dissertation Abstracts International*, *67*(3), 819A.

*Fry, M. K. (2007). The relationship of admission type and freshman seminar participation to retention and academic success. *Dissertation Abstracts International*, *67*(11), 4093.

Grubb, W. N. (2002). Learning and earning in the middle, Part I: National studies of pre-baccalaureate education. *Economics of Education Review*, *21*, 299–321.

*Hecker, M. J. (1995). Assessing the effectiveness of a developmental educational program using GPA attainment, student retention, and student involvement as measurement variables. *Dissertation Abstracts International*, *55*(9), 2739A.

Hedges, L. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486–504.

Hossler, D., Braxton, J. M., & Coopersmith, G. (1989). Understanding student college choice. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. V, pp. 231–288). New York, NY: Agathon.

Hossler, D., & Gallagher, K. S. (1987). Studying college choice: A three-phase model and the implications for policy-makers. *College and University*, *2*, 207–221.

*House, J. D., & Wohlt, V. (1991). Effect of tutoring on voluntary school withdrawal of academically underprepared minority students. *Journal of School Psychology*, *29*, 135–142.

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino students' sense of belonging. *Sociology of Education*, *70*, 324–345.

Kemple, J. J. (2008). *Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood*. New York, NY: MDRC.

Krolik, T. J. (2004). *Educational attainment of the labor force and jobless rates, 2003*. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Loiacono, C. M. (2000). Promoting freshman college student development using cognitive developmental theory presented in a deliberate psychological education-based freshman orientation program. *Dissertation Abstracts International*, *61*(6), 2186A.

Marcotte, D., Bailey, T., Borkoski, C., & Kienzl, G. S. (2005). The returns to education at community colleges: Evidence from the National Education Longitudinal Survey. *Educational Evaluation and Policy Analysis*, *27*, 157–175.

McDonough, P. M. (2004). *The school to college transition: Challenges and prospects*. Washington, DC: American Council on Education.

*McGee, S. (2004). Academic success and belonging among first-generation community college students. *Dissertation Abstracts International*, *65*(6), 2089A.

*McGregor, G. D., Jr. (2001). Creative thinking instruction for a college study skills program: A case study. *Dissertation Abstracts International*, *62*(10), 3293A.

*Milligan, J. H. (2007). The efficacy of a strengths-based study skills and strategies program for students on academic probation. *Dissertation Abstracts International*, *68*(6), 2338A.

Moss, B., & Yeaton, W. (2006). Shaping policies related to developmental education: An evaluation using the regression-discontinuity design. *Educational Evaluation and Policy Analysis*, *28*, 215–229.

National Center for Public Policy and Higher Education. (2004). *The educational pipeline: Big investment, big returns*. San Jose, CA: Author.

Nora, A., Barlow, L. & Crisp G. (2006). Examining the tangible and psychosocial benefits of financial aid with student access, engagement, and degree attainment. *American Behavioral Scientist*, *49*, 1636–1651.

Paulsen, M. B., & St. John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *Journal of Higher Education*, *73*, 189–236.

Perna, L. W. (2006). Studying college access and choice: A proposed conceptual model. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. XXI, pp. 99–157). Dordrecht, the Netherlands: Springer.

Prince, D., & Jenkins, D. (2005). *Building pathways to success for low-skill adult students: Lessons for community college policy and practice from a statewide longitudinal tracking study*. New York, NY: Community College Research Center, Teachers College, Columbia University.

Rendón, L. I., Jalomo, R. E., & Nora, A. (2000). Theoretical considerations in the study of minority student retention in higher education. In J. M. Braxton (Ed.), *Reworking the departure puzzle* (pp. 127–156). Nashville, TN: Vanderbilt University Press.

Rosenbaum, J. E. (2001). *Beyond college for all: Career paths for the forgotten half*. New York, NY: Russell Sage Foundation.

Rosenbaum, J. E., DeLuca, S., Miller, S. R., & Roy, K. (1999). Pathways into work: Short- and long-term

effects of personal and institutional ties. *Sociology of Education*, *72*, 179–196.

Rothstein, H. R., Turner, H. M., & Lavenberg, J. (2004). *Information retrieval policy brief*. Philadelphia, PA: International Campbell Collaboration.

*Salinitri, G. (2005). The effects of formal mentoring on the retention rates for first-year, low achieving students. *Canadian Journal of Education*, *28*, 853–873.

*Sanders, E. A. (2000). Project S.T.A.R.S.: A program evaluation of a freshman peer support retention program. *Dissertation Abstracts International*, *60*(12), 6425B.

Schuh, J. H. (2005). Finances and retention: Trends and potential implications. In A. Seidman (Ed.), *College student retention: Formula for student success* (pp. 277–293). Westport, CT: Greenwood.

*Scrivener, S., Bloom, D., LeBlanc, A., Paxson, C., Rouse, C. E., & Sommo, C. (2008). *A good start: Two-year effects of a freshman learning community program at Kingsborough Community College*. New York, NY: MDRC.

Shadish, W. R., & Haddock, K. C. (2009). Combining estimates of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 257–277). New York, NY: Russell Sage Foundation.

St. John, E. P., Asker, E. H., & Hu, S. (2001). The role of finances in student choice: A review of theory and research. In M. B. Paulsen & J. C. Smart (Eds.), *The finance of higher education: Theory*, *research*, *policy*, *and practice* (pp. 419–438). New York, NY: Agathon.

St. John, E. P., Cabrera, A., F. Nora, A., & Asker, E. H. (2000). Economic influences on persistence reconsidered: How can finance research inform the reconceptualization of persistence models? In J. M. Braxton (Ed.), *Reworking the departure puzzle* (pp. 29–47). Nashville, TN: Vanderbilt University Press.

St. John, E. P., Paulsen, M. B., & Carter, D. F. (2005). Diversity, college costs, and postsecondary opportunity: An examination of the financial nexus between college choice and persistence. *Journal of Higher Education*, *76*, 545–569.

St. John, E. P., Paulsen, M. B., & Starkey, J. B. (1996). The nexus between college choice and persistence. *Research in Higher Education*, *37*, 175–220.

*Stovall, M. L. (1999). Relationships between participation in a community college success course and academic performance and persistence. *Dissertation Abstracts International*, *60*(9), 3250A.

Tierney, W. (1992). An anthropological analysis of student participation in college. *Journal of Higher Education*, *63*, 603–618.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*, 89–125.

Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education*, *53*, 687–700.

Tinto, V. (1986). Theories of student departure revisited. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. II, pp. 359–384). New York, NY: Agathon.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago, IL: University of Chicago Press.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago, IL: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *Journal of Higher Education*, *68*, 599–623.

Valentine, J. C., & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: The Study Design and Implementation Assessment Device (Study DIAD). *Psychological Methods, 13,* 130–149.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need?: A primer on statistical power in meta-analysis. *Journal of Educational and Behavioral Statistics, 35,* 215–247.

Authors

JEFFREY C. VALENTINE, College of Education & Human Development, University of Louisville.

AMY S. HIRSCHY, College of Education & Human Development, University of Louisville.

CHRISTINE D. BREMER, Institute on Community Integration, University of Minnesota.

WALTER NOVILLO, College of Education & Human Development, University of Minnesota.

MARISA CASTELLANO, College of Education & Human Development, University of Louisville.

AARON BANISTER, College of Education & Human Development, University of Louisville.